

(d) delete clusters which are identified by phrases which are subset of a phrase identifying another cluster;

(e) merge clusters which contain identical documents; and

(f) display each cluster along with at least one context string from a set of documents for both the query terms and the cluster terms.

---

Please add claim 90 as follows:

---

90. (New) The method according to claim 1, further including the steps of creating a database of meta search information regarding query terms and using the database to display information relating to the query terms when a user uses those terms in a query.

---

#### REMARKS

Reconsideration of this application, as amended, is respectfully requested.

In the Office Action, the Examiner made objections to claims 84 and 85 for certain minor informalities. These informalities have been corrected pursuant to the Examiner's suggestions. These corrections are not made for any reason of patentability, but merely to correct minor typographical errors.

In the Office Action, the Examiner rejected claims 1, 16, 46, 52, and 86-89 under 35 U.S.C. § 103 (a) as being unpatentable over Redfern in view of "Text Search and Retrieval Examiner Training Manual for the Automated Patent System (APS)", hereinafter referred to as APS. In addition, the Examiner rejected claims 3-15, 18-29, 48-51, 54-57, and

80-85 under 35 U.S.C. § 103 (a) as being unpatentable over Redfern. The rejections are respectfully traversed as set forth in more detail below.

The Examiner rejected claims 1, 16, 46, and 52 on identical grounds. These grounds consist of the allegation that Redfern teaches all of the elements of these claims with the exception of locating query terms in the retrieved documents, extracting text surrounding the query terms and displaying the text surrounding the query terms. The Examiner alleges that APS discloses these features. However, as described in more detail below, the combination of Redfern and APS does not disclose, teach or suggest all of the elements of claims 1, 16, 46 and 52.

The present invention is directed to a meta search engine which forwards a query to a plurality of third party search engines, receives and processes the responses from the third party search engines, identifies documents that respond to the query including downloading the full text of the documents, locating query terms in the documents and extracting text surrounding the query terms. The text surrounding the query terms, which forms one or more context strings, and information regarding the documents is displayed. As set forth in each of claims 1, 16, 46 and 52, the information regarding the documents and the context strings surrounding the query terms are progressively displayed.

As described on page 4 of the present application, the results are returned progressively after each individual page is downloaded and analyzed, rather than after all pages are downloaded. There is no teaching or suggestion whatsoever in Redfern or APS of progressively displaying the search results. It is well established in the prior art that the display of the search results is made only after all the results have been retrieved and analyzed. In fact, in the Examiner's rejection, the description of the APS does not include a

mention of the progressive display of the search results. Thus, the Examiner has failed to make it a prima facie rejection of obviousness. Therefore, it is respectfully submitted that claims 1, 16, 46 and 52 are allowable over the hypothetical combination of Redfern and APS.

Regarding claim 86, the Examiner contends the Redfern teaches clustering as recited in claim 86. Claim 86 includes the step of clustering the documents based on analysis of the full text of each document that was retrieved in response to the query. In addition, the method identifies co-occurring phrases and words and conjunctions thereof found in these result documents and displays the documents arranged according to these clusters. Document clustering is described in the specification on pages 20-22. Clustering of claim 86 provides a significant aid to information discovery by identifying for the user what topics are found most often in all of the documents that are retrieved in response to the query. This will allow the user to further refine the query in order to investigate one of these topics. There is no teaching or suggestion whatsoever in Redfern of clustering the documents retrieved in response to the query. In contrast, Redfern is directed to a detailed method of parsing and refining natural language search queries. All of the sections cited by the Examiner in connection with the rejection of claim 86 are directed to simplifying the query, not to clustering the documents retrieved in result to the query and for displaying information regarding the documents arranged by the clusters. Thus, claim 86 is allowable over the hypothetical combination of Redfern and APS.

The Examiner's rejection of claim 87 is similarly defective in that claim 87 is directed to displaying suggested additional query terms for expanding the query based on the terms in the documents identified in response to the query. Again, Redfern is directed to revising and simplifying query terms based on the terms themselves, not based on documents

that are retrieved that contained query terms. Thus, claim 87 is allowable over the hypothetical combination of Redfern and APS.

With respect to claim 88, the Examiner alleges that Redfern teaches receiving a query and transferring the query from a form of a question into a form of an answer prior to forwarding the query to the plurality of third party search engines. The Examiner cites column 15, lines 42-49 wherein a query in the form of a question is shown, namely, “where do monarch butterflies spend the winter?” The example in Redfern shows that for one search engine, the query would be parsed into the form “monarch + butterflies +spend + winter. This form of parsing takes the keywords from the query and repeats them in the order they appear in the query. There is no teaching or suggestion to transform the query into a form of an answer. The Examiner’s attention is directed to the present specification beginning on the bottom of page 23 to page 26. One example described therein is the query “what does NASDAQ stand for?” which is transformed in accordance with the present invention to “NASDAQ is an abbreviation”. Another example of a transformed query would be “NASDAQ means”. In contrast, Redfern does not teach transforming queries into a form of an answer, but rather merely parsing the queries into the format requirements for the various search engines. Therefore, it is submitted that claim 88 is allowable over the hypothetical combination of Redfern and APS.

Claim 89 includes a step of displaying an indication of how close the query terms are to each other in the documents in addition to displaying the information regarding the documents and the context strings surrounding the query terms for each document. The section of Redfern cited by the Examiner discloses a process for ranking the information sources. A system that ranks sources is not one that ranks the documents based on how close

the query terms are in the documents. In contrast, Redfern teaches ranking which search engine provided the most relevant documents. Therefore, claim 89 is allowable over the hypothetical combination of Redfern and APS.

Claim 4 is directed to identifying and displaying a list of documents identified in response to the query which do not contain any of the query terms. It is possible for various reasons that documents identified in response to the query search do not actually contain any of the query terms. Claim 4 is directed to the display of a separate page that lists these documents. There is no teaching in Redfern of this feature. The section of Redfern, namely, column 16, lines 1-11, cited by the Examiner does not disclose such a feature. Therefore, claim 4 is allowable over Redfern.

Claim 5 is allowable over Redfern for the same reasons described above with respect to claim 86.

Claim 6 is directed to storing documents matching a query so that a query can be repeated and only showing documents which are new or have been modified since the last query or given time. The sections of Redfern cited by the Examiner are merely directed to the display of hits for a query. There is absolutely no disclosure of the feature of updating a query so that only new documents that have come on line subsequent to the previous query are displayed. Thus, claim 6 is clearly allowable over Redfern.

Claim 8 and new claim 90 are directed to the creation of a database of meta search information regarding query terms for the purpose of displaying information relating to the query terms when a user uses those terms in a query. There is no teaching or disclosure whatsoever in Redfern of creating a database of meta search information regarding query terms and displaying a page related to the query terms set forth in new claim 90 and claim 8.

Moreover, the Examiner has provided absolutely no evidence to support the allegation that it would have been obvious to one of ordinary skill in the art to incorporate the feature of claims 90 and 8 in a meta search engine. Thus, claims 90 and 8 are allowable over Redfern.

Claim 9 is directed to storing and using information regarding the particular documents requested by the user in response to a query by remembering the most commonly requested document for a given query and presenting this document first in response to the same query in the future. The section of Redfern cited by the Examiner, namely, column 1, lines 21-65 merely describes the prior art of presenting a search query and returning documents containing the query terms. Remembering commonly requested documents for future queries is clearly not discussed in Redfern. Therefore, claim 9 is allowable over Redfern.

Claim 10 is directed to analyzing the number of documents which have been found as a function of the number of third party search engines queried, and computing the estimated size of the third party search engines and the estimated size of the document base which the third party search engines index. The sections cited by the Examiner in Redfern are directed to the methods of ranking the sources from which the documents retrieved are obtained, to determine which source has the most relevant documents. There is no teaching or suggestion whatsoever in Redfern of estimating the size of the third party search engine and the size of the document base which the third party search engine index. Thus, claim 10 is clearly allowable over Redfern.

Claim 12 is directed to detecting and displaying duplicate documents by identifying duplicate context strings. First, the sections cited by the Examiner do not even deal with duplicates. Secondly, it is known in the prior art that duplicate pages are

determined based on title of the document or possibly the header or footer. The present invention is directed to determining that a document is a duplicate by looking at the actual text in the document and comparing the context string that includes the query term and comparing these context strings. There is no teaching or suggestion of this feature in Redfern. Therefore, claim 12 is allowable over Redfern.

Claim 14 is allowable for the same reasons set forth above with respect to claim 87.

Claim 15 is directed to re-ranking documents according to the number and proximity between query terms and redisplaying the information regarding the documents according to the ranking. The sections cited by the Examiner in Redfern mention ranking, but there is no teaching or suggestion of a re-ranking of the documents according to the number of and proximity between query terms and nor is there a redisplay of the information regarding the documents according to this re-ranking. As noted above, Redfern teaches ranking the sources, but there is no ranking of each document displayed from each source. In contrast to Redfern which rank document sources, claim 15 is directed to ranking the documents themselves. Therefore, claim 15 is allowable over Redfern.

Claim 80 is allowable for the same reasons set forth with respect to claim 88.

Claim 81 is directed to displaying an indication of how close the query terms are to each other in the documents and the information regarding each document. The section cited in Redfern by the Examiner is directed to a ranking of the sources, which is merely based on the number of times the query terms appear. There is no teaching or suggestion in Redfern of a ranking system based on how close the query terms are to each other. Therefore, claim 81 is allowable over Redfern.

Claim 82 is directed to displaying the search results in five different ways, namely, (a) a predetermined number of documents ranked using term proximity information; (b) documents which contain less than all the query terms; (c) documents that contain one of the query terms; (d) documents that contain duplicate context strings ; and (e) documents that could not be downloaded. None of the sections cited by the Examiner in Redfern teach any of the five methods for displaying documents set forth in claim 82. Moreover, there is no teaching in Redfern whatsoever of any one display result displaying the documents in these five different ways. Thus, claim 82 is clearly allowable over Redfern.

Claim 83 is allowable for the same reasons set forth with respect to claim 87 above.

Claim 84 is directed to displaying summary information regarding the documents found and processed, the summary information being separately identified for each search engine. An example of the summary information is shown in Fig. 6 of the present application. As can be seen in Fig. 6, each of the search engines is listed together with the total number of responses, the number of documents retrieved, the number of documents processed and the number of documents that were duplicates. The sections cited by the Examiner in Redfern do not disclose the summary information for each of the search engines. Appendix G lists the actual documents retrieved together with an excerpt from a document. The information shown in Exhibit G is not summary information and, therefore, claim 84 is allowable over Redfern.

Claim 85 is directed to a specific clustering algorithm disclosed in the present specification. There is no teaching or disclosure whatsoever in Redfern of a clustering

algorithm and there is clearly no disclosure of the specific algorithm set forth in claim 85.

Therefore, claim 85 is clearly allowable over Redfern.

Claims 18-29, 48-50 and 54-57 are allowable for the reasons described above with respect to the claims relied upon by the Examiner for the basis of the rejections, and therefore, there is no need to repeat those arguments here.

Attached hereto is a marked-up version of the changes made to claims by the current amendment. The attached page is captioned **“Version with Markings to Show Changes Made”**.

Based on the above, it is respectfully submitted that all of the claims in the application contain allowable subject matter. Therefore, applicants respectfully request that a timely Notice of Allowance be issued in this case.

Respectfully submitted,



Paul J. Esatto, Jr.

Registration No. 30,749

Scully, Scott, Murphy & Presser  
400 Garden City Plaza  
Garden City, New York 11530  
(516) 742-4343  
PJE:ae

09/113/751

**“VERSION WITH MARKINGS TO SHOW CHANGES MADE”.**

Claims 8, 9, 10, 84 and 85 have been amended as follows:

8. (Amended) A method according to Claim [1] 90, [further including] wherein the steps of creating and using a database of meta-information regarding query terms, [e.g.] includes storing a list of movie titles, recognizing when the user enters a query containing a movie title, and taking a special action such as referring the user to the review of the movie at a specific movie review site.

9. (Amended) A method according to Claim 1, further including the step of storing and using information regarding the particular documents requested by a user in response to a query[, e.g.] by remembering the most commonly requested document for a given query and presenting this document first in response to the same query in the future.

10. (Amended) A method according to Claim 1, further including the steps of analyzing the number of documents which [would] have been found as a function of the number of third party search engines queried, and computing the estimated size of the third party search engines and the estimated size of the document base which the third party search engines index.

84. (Amended) The method of Claim [84] 82, further including the step of displaying summary information regarding the documents found and processed, the summary information being separately identified for each search engine.

85. (Amended) The method of Claim 5, wherein the step of clustering comprises for each processed document the steps of:

- (f) for  $n = 1$  to MaximumPhraseLength, for each set of successive  $n$  words, if this combination of words has not already appeared in this document, then add the set to a hash table for this document and a hash table for all documents;
- (g) for  $n = \text{MaximumPhraseLength}$  to 1, find the most common phrases of length  $n$ , to a maximum of MaxN phrases, which occurred more than MinN times, and add these phrases to the set of clusters;
- (h) find the most common combination of two clusters [form] from the previous step, to a maximum of maxC combinations, for which the combination occurred in individual documents at least MinC times;
- (i) delete clusters which are identified by phrases which are subset of a phrase identifying another cluster;
- (j) merge clusters which contain identical documents; and
- (f) display each cluster along with at least one context string [form] from a set of documents for both the query terms and the cluster terms.